Accelerated search for $BaTiO_3$-based piezoelectrics with vertical morphotropic phase boundary using Bayesian learning

Dezhen Xue[a,b], Prasanna V. Balachandran[a], Ruihao Yuan[b], Tao Hu[c], Xiaoning Qian[c], Edward R. Dougherty[c], and Turab Lookman[a,1]

[a]Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545; [b]State Key Laboratory for Mechanical Behavior of Materials, Xi'an Jiaotong University, Xi'an 710049, China; and [c]Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843

Edited by David Vanderbilt, Rutgers, The State University of New Jersey, Piscataway, NJ, and approved October 4, 2016 (received for review May 9, 2016)

An outstanding challenge in the nascent field of materials informatics is to incorporate materials knowledge in a robust Bayesian approach to guide the discovery of new materials. Utilizing inputs from known phase diagrams, features or material descriptors that are known to affect the ferroelectric response, and Landau–Devonshire theory, we demonstrate our approach for $BaTiO_3$-based piezoelectrics with the desired target of a vertical morphotropic phase boundary. We predict, synthesize, and characterize a solid solution, $(Ba_{0.5}Ca_{0.5})TiO_3$-$Ba(Ti_{0.7}Zr_{0.3})O_3$, with piezoelectric properties that show better temperature reliability than other $BaTiO_3$-based piezoelectrics in our initial training data.

piezoelectric materials | materials informatics | Bayesian learning | morphotropic phase boundary | Pb-free materials

Accelerating the process of materials design and discovery is an emerging theme in materials science (1). The emphasis has so far largely been on screening databases or using data-driven approaches that infer predictions directly from the data, be it from high-throughput calculations or experimental measurements (2–6). However, a distinguishing aspect of materials science is that in addition to data there exists a substantial body of knowledge in the form of phenomenological models and physical theories that could be used to constrain the inference models. Hence, a key challenge in materials informatics is to incorporate knowledge to make predictions that are more robust than would be possible by using data alone. Although such knowledge is used in choosing features or descriptors for materials informatics (7–9), it has seldom been used to encode prior information in the form of probability distributions and uncertainties for predicting novel materials with desired properties. Bayesian inference, which permits integration of prior knowledge or beliefs with the observed data, has shown considerable promise in cancer genomics (10) using metabolic pathway information, and in systems biology (11), but has been little explored in materials science. Our objective is to combine empirical data and materials knowledge within a Bayesian approach coupled to the results of Landau–Devonshire theory (12, 13) to design better $BaTiO_3$-based lead-free piezoelectrics.

Piezoelectric materials, such as the solid solutions of $BaTiO_3$, are best suited for exploring Bayesian inference methods because historically they are well modeled by Landau–Devonshire theory (12–14) and equations exist for describing some of the key characteristics that determine the functional response, such as the morphotropic phase boundary (MPB) (15, 16). These equations serve as "constraints" that encode prior knowledge within our Bayesian formalism. Furthermore, $BaTiO_3$-based solid solutions represent an important class of potential substitutes for Pb-based materials, which suffer from environmental concerns. Akin to the Pb-based piezoelectrics, MPBs can be established in $BaTiO_3$-based solid solutions that enable polarization and structural instabilities giving rise to a large electromechanical response. For example, Liu and Ren (17) synthesized and characterized the solid solution $x(Ba_{0.7}Ca_{0.3})TiO_3-Ba(Zr_{0.2}Ti_{0.8})O_3$ and reported a piezoelectric response, d_{33} (the longitudinal piezoelectric strain coefficient), of the order of 600 pC/N at the MPB, which is even greater than that for polycrystalline lead zirconate titanate. However, it suffers from inferior temperature reliability (18), that is, for a fixed MPB composition d_{33} varies substantially with temperature. This inferior temperature reliability is a characteristic feature of other Pb-free piezoelectric systems as well, for example, $Bi_{0.5}Na_{0.5}TiO_3$ and $K_{0.5}Na_{0.5}NbO_3$, and is reflected in the tilted or curved MPB in the temperature vs. composition phase diagram (19). As shown in Fig. 1*A*, a vertical MPB ensures that the piezoelectric material is always in the vicinity of the MPB region as a function of temperature, but for a curved or tilted MPB, as the temperature deviates from the MPB line, the system is further removed from structural instabilities and the piezoelectric properties are severely affected, as shown schematically in Fig. 1*B*. Therefore, our design target is to find novel $BaTiO_3$-based lead-free piezoelectrics with a vertical MPB (i.e., from Fig. 1*B* to Fig. 1*A*) so that the piezoelectric response is not compromised over the entire operating temperature range.

One of the common approaches to designing MPBs in solid solutions is to combine a tetragonal (T) end compound [which undergoes a cubic (C)-to-T transition] with a rhombohedral (R) end compound (C-to-R transition) to form a pseudobinary solid solution (17). This approach has led to a number of $BaTiO_3$-based solid solutions, such as $x(Ba_{0.7}Ca_{0.3})TiO_3-Ba(Zr_{0.2}Ti_{0.8})O_3$, $x(Ba_{0.7}Ca_{0.3})TiO_3-Ba(Sn_{0.12}Ti_{0.88})O_3$ and $x(Ba_{0.7}Ca_{0.3})TiO_3-Ba(Hf_{0.2}Ti_{0.8})O_3$ (17, 20, 21), which afford ease of synthesis but are characterized by an MPB with significant curvature. The choice of dopant added to

Significance

Learning from data to accelerate the discovery of new materials is an outstanding challenge in materials science. However, in addition to data, a distinguishing aspect of materials science is that there exists a substantial body of knowledge in the form of phenomenological models and physical theories. We combine informatics and materials knowledge using results from Landau–Devonshire theory to guide experiments in the design of lead-free piezoelectrics with desired properties.

Author contributions: D.X., P.V.B., and T.L. designed research; D.X., P.V.B., R.Y., T.H., X.Q., E.R.D., and T.L. performed research; D.X., P.V.B., T.H., X.Q., E.R.D., and T.L. analyzed data; and D.X., P.V.B., T.H., X.Q., E.R.D., and T.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. Email: txl@lanl.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1607412113/-/DCSupplemental.

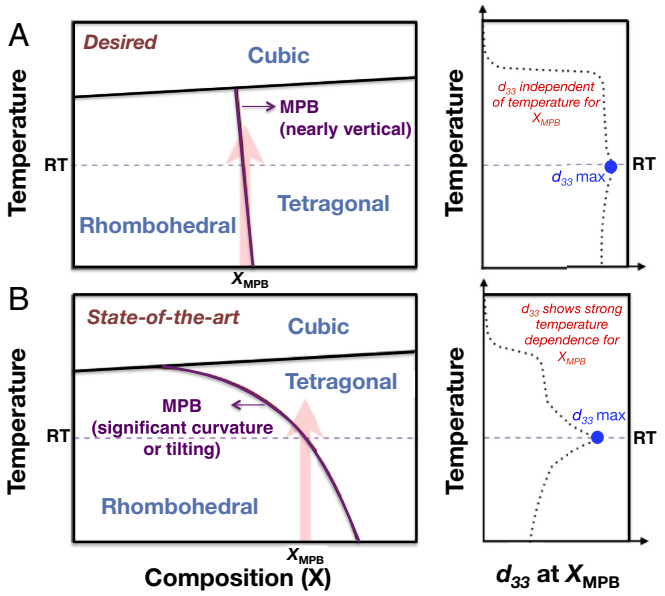

Fig. 1. Importance of vertical MPB for superior piezoelectric performance in the composition–temperature phase diagram for ferroelectric systems. A vertical MPB, as schematically shown in *A*, provides temperature-independent d_{33} piezoelectric property (*Right*) for the MPB composition (X_{MPB}), which is desirable for practical applications. In contrast, a tilted MPB (as shown in *B*) gives rise to highly temperature-sensitive d_{33} property (*Right*), which is undesirable. Notice that in *A* with the vertical MPB, d_{33} max is sustained for a larger temperature interval for the X_{MPB} composition. However, in *B* for the composition X_{MPB}, d_{33} max occur only at room temperature (RT), because only at this temperature X_{MPB} resides at the MPB. All known $BaTiO_3$-based piezoelectric materials exhibit tilted MPB as shown in *B*. Our objective is to discover new chemical compositions that may have the desired vertical MPB in $BaTiO_3$-based solid solutions.

$BaTiO_3$, as well as their relative concentrations, can affect the MPB shape and curvature. Hence, an outstanding challenge is to predict optimal chemical compositions of dopants at the T and R ends that will give rise to a vertical MPB. This is a nontrivial task because it is an inverse problem with a vast search space; it can be as large as having to experimentally determine over 1,200 phase diagrams, and each phase diagram requires the synthesis of several (6 to 14) different compositions. Thus, mapping out the phase diagrams is very time-consuming and has been done for only a handful of compositions. Moreover, the features that control the MPB in the phase diagram are also not known. There are also no a priori models in the literature for the predictive design of piezoelectrics with targeted MPB curvature, and trial and error and intuition have been the primary means by which experimentalists have tried to search for new compositions.

Although high-throughput ab initio calculations have been used for piezoelectric materials design (22), these efforts are not suitable for exploring a vast search space involving solid solutions and predicting temperature dependence of MPB compositions. Whereas the emphasis in materials informatics over the last few years has centered on discovering new materials with targeted properties, the use of information sciences in materials is not new. Chelikowsky and Phillips (23) studied the classic problem of classifying AB *sp*-bonded octet solids with sixfold coordinated rock salt or fourfold coordinated zinc blende/wurtzite in the 1970s. They constructed 2D maps to classify the AB compounds, and this problem has been revisited recently from a statistical learning perspective (8, 24, 25). A range of materials regression problems, from melting temperatures of solids (26) to dielectric constants and band gaps of polymers (27), have also been studied using machine learning tools such as kernel ridge and support vector regression. Bayesian-based Gaussian process models have also been applied to predict thermal conductivity in solids (28). However, these approaches have largely been data-driven with materials knowledge used to construct features. Although some of the recent adaptive strategies can potentially overcome certain shortcomings (6, 29), a principled approach requires integrating prior information with available data within a Bayesian framework. Few attempts, if any, in the materials literature incorporate theory within a Bayesian formalism to constrain the model outcomes. This becomes important for materials design problems involving small data sizes, especially when the objective is to make predictions in a relatively large search space (such as those discussed here).

Here, we develop a Bayesian approach for piezoelectric materials design that uses the experimental phase diagrams and Landau–Devonshire theory of ferroelectric phase transitions, in conjunction with the inference model, to predict the composition $x(Ba_{0.5}Ca_{0.5})TiO_3-Ba(Zr_{0.30}Ti_{0.70})O_3$. Synthesis reveals that this solid solution has, indeed, 15% less MPB curvature and improved temperature-insensitive d_{33} response relative to the best-known material in our training dataset. Further iteration via feedback did not improve the predictions. Our work demonstrates how functional relationships from Landau theory, when integrated into a Bayesian learning framework, can guide the design of new materials with targeted response.

Results

Our Bayesian approach for materials design involves three key elements, namely, data, prior knowledge, and learning, for guiding the experimental synthesis efforts.

Data.

Property. Our approach involves building a regression model, $y = f(z) \pm \sigma$, which enables prediction of y with associated uncertainties (σ) from features (z). We learn f from known experimental data and, in turn, use it to predict y with σ for all compositions in the unexplored or "virtual" dataset. Because the MPB is a function of temperature τ (in °C) and concentration x (Fig. 1), we perform a quadratic fit, $\tau(x) = A_1x^2 + A_2x + C$, in which we identify three coordinates in the (x, τ) phase diagram: (*i*) $(x_1, 25)$, (*ii*) $(x_1, -75)$, and (*iii*) $(x_2, -75)$, where 25 and -75 are temperatures in °C and the coordinates $(x_1, 25)$ and $(x_2, -75)$ should lie on the MPB curve. The difference between x_2 and x_1 (the y coordinates of which correspond to -75°C) is denoted by dx and represents the degree of MPB curvature (*SI Appendix*, Note 1 and Fig. S1). For each of the 19 phase diagrams we estimated dx. Our objective is to identify $BaTiO_3$-based solid solutions with smaller dx than in the training set.

Features. Because the "verticality" of the MPB is closely related to the choice of two end-member compositions, we used features that couple the atomic, crystal chemistry and electronic structure properties of T and R ends for regression (30–35). Because polarization and strain are the primary and secondary order parameters, respectively, of the electromechanically coupled piezoelectric materials, we chose the following features: (*i*) the unit cell volume difference between the T and R ends, $\Delta Vol = V_T - V_R$, where V_T and V_R are the unit cell volumes of the T and R ends; (*ii*) t_f, defined by $t_f = \frac{R_T^A}{R_R^B}$, where R_T^A is the weighted average ionic radii of the A site ion at the T end and R_R^B is the weighted average ionic radii of the B site ion at the R end; (*iii*) the ionic displacements for the T (δD_T) and R (δD_R) ends, respectively, taken from the work of Balachandran et al. (33);

(*iv*) the ratio of the effective nuclear charge between the A site ion at the T end and B-site ion at the R end, r_{enc}, defined by $r_{enc} = \frac{A_{enc_T}}{B_{enc_R}}$; and (*v*) the ratio of electronegativities between the A site ion at the T end and B-site ion at the R end, r_{en}, defined by $r_{en} = \frac{A_{en_T}}{B_{en_R}}$. Each end member composition is uniquely described as a weighted fraction of each of these features.

Chemical design space. Our strategy is to combine Ca^{2+}-doped $BaTiO_3$ (T end) with Sn^{4+}- or Zr^{4+}-doped $BaTiO_3$ (R end) to give rise to a solid solution with an MPB whose curvature can be tuned by varying the relative concentrations of Ca^{2+}, Sn^{4+}, or Zr^{4+} cations. We considered two pseudobinary ceramic systems: (*i*) $x(Ba_{1-m}Ca_m)TiO_3-Ba(Zr_nTi_{1-n})O_3$ (xBCT–BZT) and (*ii*) $x(Ba_{1-m}Ca_m)TiO_3-Ba(Sn_nTi_{1-n})O_3$ (xBCT–BTS) and our objective is to find an optimal m and n such that the curvature of the MPB in the ceramic is minimized. For Ca^{2+}, the lower limit for m is set at 18% to maintain the pure C-to-T transition and the upper limit is at 50%, above which secondary phases begin to appear, that is, $18\% \leq m \leq 50\%$. Similarly, for Sn^{4+} and Zr^{4+}, the range for n is $10\% \leq n \leq 30\%$ and $15\% \leq n \leq 30\%$, respectively, to ensure a pure C-to-R transition. The resolution of our experiments allowed control of m and n compositions to 1%, giving rise to a total chemical space of 1,222 potential phase diagrams. Of these 19 are already known or reported, which served as our training data (*SI Appendix*, Fig. S2 and Table S1), and none of the 19 had a vertical MPB resembling that of Pb-based piezoelectrics. The remaining phase diagrams formed a virtual space consisting of 1,203 potential samples to be explored (*SI Appendix*, Note 2). At least 6 to 14 unique x compositions (*SI Appendix*, Fig. S2) need to be synthesized for one phase diagram for given m and n. This significantly expands our total number of experiments to 7,800 to 18,200 compositions to be explored, representing a vast chemical and complex space, which becomes even larger if the compositions are controlled to a higher resolution.

Prior Knowledge. Our Bayesian design strategy is illustrated in Fig. 2.

For over 50 years, Landau–Devonshire theory has been widely used to reproduce phase diagrams for many piezoelectrics and study their performance at the MPB. The $m3m$ cubic symmetry of the paraelectric phase allows transformations to seven possible ferroelectric monodomain phases that are characterized by different polarization vectors, $\mathbf{p} = \mathbf{n}p$, where $\mathbf{n} = \{n_1, n_2, n_3\}$ is a unit vector in the direction of polarization and p is its absolute value (14). Thus, the free energy, g, of the ferroelectric system can be described by a Landau polynomial that depends on the modulus of the polarization vector (p) and the polarization direction ($\mathbf{n}$), which is given by

$$\begin{aligned} g(\tau, \bar{n}, p) &= \frac{1}{2}\alpha_1({n_1}^2 + {n_2}^2 + {n_3}^2)p^2 + \frac{1}{4}\beta_1({n_1}^2 + {n_2}^2 + {n_3}^2)^2 p^4 \\ &+ \frac{1}{4}\beta_2({n_1}^4 + {n_2}^4 + {n_3}^4)p^4 + \frac{1}{6}\gamma_1({n_1}^2 + {n_2}^2 + {n_3}^2)^3 p^6 \\ &+ \frac{1}{6}\gamma_2({n_1}^6 + {n_2}^6 + {n_3}^6)p^6 + \frac{1}{6}\gamma_3({n_1}^2{n_2}^2{n_3}^2)p^6, \end{aligned} \quad [1]$$

where α_i, β_i, and γ_i are coefficients that are materials-dependent and determined from experiments. Coefficient $\beta_2(\tau, x)$ depends on the temperature (τ) and composition (x).

The polarization direction determines the stable ferroelectric phase so that for tetragonal phase $\mathbf{n} = (\mathbf{0}, \mathbf{0}, \mathbf{1})$ and for rhombohedral phase $\mathbf{n} = (\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}})$. The MPB is a phase boundary between the T and R phase, where the two phases coexist and

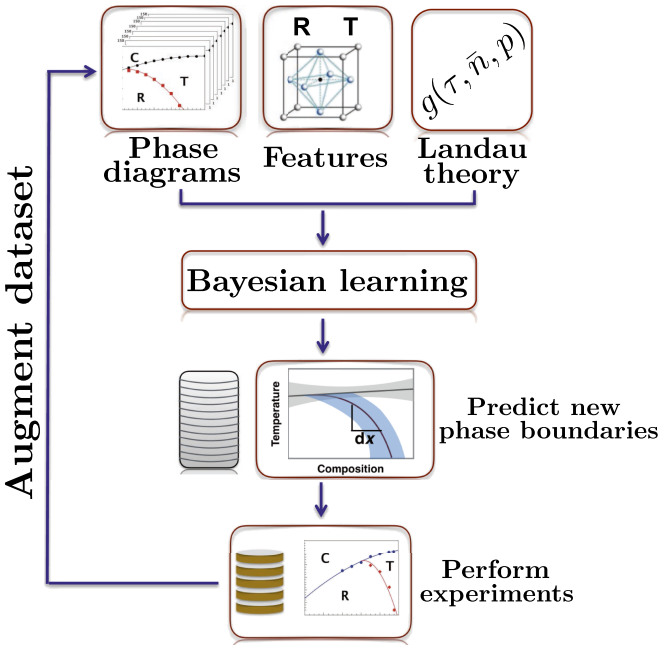

Fig. 2. Bayesian learning for materials design. An initial experimental dataset of 19 phase diagrams and features that are known to affect the MPB serve as input to learning. The Landau functional form for MPB serves as prior knowledge to constraint the model space. The resulting Bayesian linear regression model relates phase boundaries (τ_{MPB} and τ_{PF}) to materials features, where τ_{MPB} and τ_{PF} are MPB and paraelectric-to-ferroelectric phase boundaries, respectively. The model is trained and cross-validated with the initial data. The trained model with updated prior and posterior distributions is applied to a dataset of unexplored phase diagrams to predict *dx*. The "best" candidate with smallest *dx* is chosen for synthesis and characterization. We then augment the initial dataset with the newly measured *dx* to update the inference model. The loop is repeated until the material with the desired response is discovered.

are degenerate in energy, thus at the MPB (where $\tau = \tau_{MPB}$ and $x = x_{MPB}$), we require $g_T = g_R$. This gives

$$\beta_2(\tau_{MPB}, x_{MPB}) + \frac{1}{27}(24\gamma_2 - \gamma_3){p_{eq}}^2(\tau_{MPB}, x_{MPB}) = 0, \quad [2]$$

for the MPB, where p_{eq} is the equilibrium value of the polarization and $\beta_2(\tau, x) = \xi_0(x_{MPB} - x_0) + \frac{1}{2}\xi_0'(x_{MPB} - x_0)^2$. It is also known that p^2 is proportional to the transition temperature τ_C (36), that is, $p^2(\tau, x) = \rho(\tau - \tau_C(x))$ (36), where ρ is a constant. The MPB curve is then given by

$$\tau_{MPB} = \tau_C(x) - \phi(\xi_0(x_{MPB} - x_0) + \xi_0'(x_{MPB} - x_0)^2), \quad [3]$$

where $\tau_C(x)$ is the concentration dependence of Curie temperature, x_0 is the composition at the triple point, $\phi = \frac{27}{27\xi + \rho(24\gamma_2 - \gamma_3)}$, and ξ_0, ξ_0' are coefficients for $\beta_2(\tau, x)$. We now have a functional form (Eq. **3**) for the MPB curves in our phase diagrams, from which dx can be calculated.

Learning. The Landau functional form serves as the input to our Bayesian regression approach to directly predict both MPB and paraelectric–ferroelectric (PF) phase boundaries based on the aforementioned six features: $(\Delta Vol, tf, \delta D_T, \delta D_R, r_{enc}, r_{en})$. We consider quadratic forms for both the phase boundaries (MPB and PF): $\tau_{MPB}(x) = \omega_1 x^2 + \omega_2 x + C_1$ and $\tau_{PF}(x) = \eta_1 x^2 + \eta_2 x^2 + C_2$, where the coefficients $(\omega_1, \omega_2, C_1, \eta_1, \eta_2, C_2)$ are assumed to be dependent on the features. For example, for the MPB, $\omega_1 = \boldsymbol{f}^T\boldsymbol{\zeta}_{\omega_1}$ with the corresponding regression

coefficient vector ζ_{ω_1}. The feature vector $\boldsymbol{f}$ is constructed as a column vector using the six features and their products, for example, r_{en}, r_{en}^2 and $r_{en}r_{enc}$. This results in a feature vector with 27 components whose transpose is denoted as $\boldsymbol{f}^T$. From different concentrations of x, the response τ_{MPB} on the MPB can be estimated as $\tau_{MPB} = \boldsymbol{f}^T\zeta_{\omega_1}x^2 + \boldsymbol{f}^T\zeta_{\omega_2}x + \boldsymbol{f}^T\zeta_{C_1} + \sigma_{MPB} = [\boldsymbol{f}^Tx^2, \boldsymbol{f}^Tx, \boldsymbol{f}^T]\boldsymbol{\zeta} + \sigma_{MPB}$, where $\boldsymbol{\zeta} = [\zeta_{\omega_1}; \zeta_{\omega_2}; \zeta_{C_1}]$ is the regression vector with 81 coefficients, and σ_{MPB} denotes the i.i.d. Gaussian noise. Similarly, the response on the PF boundary is also estimated.

We assume a Gaussian prior distribution for $\boldsymbol{\zeta}$, which is further constrained by the domain knowledge that (*i*) the two phase boundaries intersect within a certain range of x and (*ii*) the MPB is concave ($\omega_1 < 0$). Such linear constraints result in a truncated Gaussian distribution for $\boldsymbol{\zeta}$, from which a posterior distribution is learned using the initial training set of 19 phase diagrams. The posterior is also a truncated Gaussian distribution, which does not allow for an analytical solution for the distribution of response (e.g., y_{MPB}) on the phase boundaries. Instead, we draw random samples from the learned posterior distribution of $\boldsymbol{\zeta}$ and then use the resulting samples to predict the ensemble of phase boundaries. We then calculate the change of x, denoted by dx, along each MPB when the temperature decreases by 100 K from room temperature (298 K). From the resulting ensemble of dx, we calculate its mean and uncertainties. The measured vs. predicted dx (along with uncertainties) from the training data are shown in *SI Appendix*, Fig. S3. Our Bayesian linear regression model performs reasonably well on the training data. Based on this model, we then use the randomly drawn samples of $\boldsymbol{\zeta}$ to predict the ensemble of phase boundaries and dx for each candidate material in the "virtual dataset" (see *SI Appendix*, Note 3 for more details). In addition, we filtered out materials whose predicted triple points are below 298 K.

The combination of m and n with the smallest predicted mean value dx is chosen as the best candidate for experimental validation. For the BTS system, the best candidate is BTS-m50-n30. However, its predicted triple point is below 298 K and the predicted dx is relatively larger than that for compound BZT-m50-n30, the best candidate identified for the BZT system. Thus, the Bayesian learning recommended BZT-m50-n30 for experimental synthesis.

We also used a data-driven approach (6, 29) that does not use any a priori knowledge of the functional forms of the phase boundaries. Additional details are given in *SI Appendix*, Notes 4 and 5 and Table S2. We note that they also pointed to choosing the BZT-m50-n30 compound for the next experiments.

Experimental Synthesis and Feedback. We synthesized BZT-m50-n30 using the conventional powder processing route (*Methods*). The transition temperatures were determined by dielectric permittivity (ϵ) vs. temperature (T) curves for a total of nine different compositions, x, from which we constructed the phase diagram. In Fig. 3*A* we show the experimental phase diagram (dots) and overlay it with the predicted phase boundaries from the Bayesian linear regression model (shown as lines). We find excellent agreement between the two curves. The experimental data from our newly reported phase diagram for BZT-m50-n30 was then augmented to the training set and we retrained the models, now using 20 data points to perform a second iteration of our loop. We found that our Bayesian learning picked the same compound BZT-m50-n30 from the given virtual set. The estimated dx from the updated model does not change much from the previous estimate but the uncertainties are significantly reduced (Fig. 3*B*). We conclude that BZT-m50-n30 represents the optimal material available to us among similar piezoelectrics within the $BaTiO_3$ family of compounds explored in this work. The dx of 0.1458 for BZT-m50-n30 is a 15% improvement over the best-known training sample BZT-m10-n20.

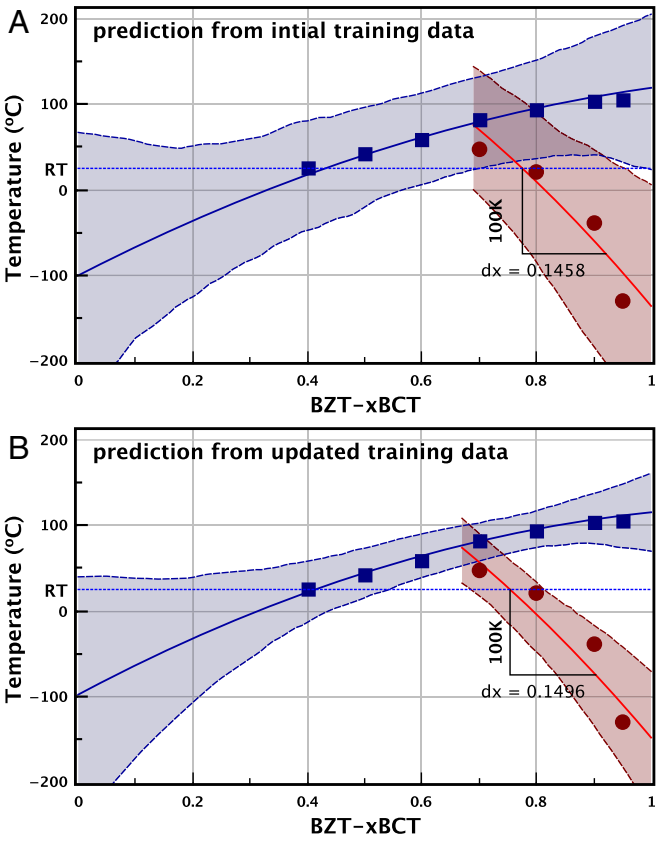

Fig. 3. (*A*) Predicted (solid lines) vs. experimental (dots) phase diagram for BZT-m50-n30 from Bayesian linear regression. The blue and red solid lines show the mean phase boundaries, and the blue and red dashed lines mark the 95% confidence intervals. (*B*) Updated Bayesian linear regression model after augmenting the experimental BZT-m50-n30 data. Notice the reduction in uncertainties (dashed lines) for the updated model.

In Fig. 4*A* we compare the best-known compound with our newly found system. Having designed BZT-m50-n30 with a relatively vertical MPB, we measured its temperature-dependent d_{33} response for the MPB composition. Results comparing BZT-m50-n30 with the most widely studied BZT-m10-n20 are shown in Fig. 4*B*. We find that the d_{33} for BZT-m50-n30 is relatively less sensitive to temperature compared with BZT-m10-n20; however, its absolute value of 85 pC/N is much lower than that for BZT-m10-n20 (500 to 600 pC/N) (Fig. 4*B*, *Inset*). Even though we achieved our design target for better MPB verticality, it was at the expense of a relatively small d_{33}. To this end, our work points to the need for developing strategies that can search for optimal compounds that simultaneously satisfy two objectives: (*i*) a vertical phase boundary and (*ii*) large d_{33} and polarization response.

Density Functional Theory. We performed density functional theory (DFT) calculations within the generalized gradient approximation (37) as implemented in Quantum ESPRESSO (38) to glean insights into why our BZT-m50-n30 compound has a smaller d_{33} and polarization compared with the $Ba(Zr_{0.2}Ti_{0.8})O_3$−$x(Ba_{0.7}Ca_{0.3})TiO_3$ system (17) (*Methods*). The solid solutions were treated within the virtual crystal approximation (39), which has been shown to reproduce, with reasonable accuracy, the ground-state properties of ferroelectric solid solutions (40). We chose two systems, $Ba(Zr_{0.2}Ti_{0.8})O_3$−$x(Ba_{0.7}Ca_{0.3})TiO_3$ (17) and BZT-m50-n30 predicted in this work, whose MPB compositions correspond to $Ba_{0.85}Ca_{0.15}Ti_{0.9}Zr_{0.1}O_3$ (BCTZ1) and $Ba_{0.9}Ca_{0.1}Ti_{0.74}Zr_{0.26}O_3$ (BCTZ2), respectively.

We computed the total energy difference for different crystal symmetries and spontaneous polarization from the Berry phase (41) method.

Our results show that in BCTZ1 composition the total energy differences (ΔE) between cubic to tetragonal, orthorhombic, and rhombohedral phases are -17, -20 and -22 meV/f.u., respectively (Table 1). The negative sign indicates that the low-temperature ferroelectric phases have lower energy than that for the high-temperature cubic phase, in agreement with the experimental phase diagram. The spontaneous polarization ΔP_s for tetragonal, orthorhombic, and rhombohedral phases are calculated as 35, 38.4, and 38.2 $\mu C/cm^2$, respectively. However, in our newly found MPB composition BCTZ2, the ΔE for cubic to tetragonal, orthorhombic, and rhombohedral phases are -0.6, -0.3 and -0.5 meV/f.u., respectively. In addition, ΔP_s for tetragonal, orthorhombic, and rhombohedral phases are 0.07, 0.05, and 0.06 $\mu C/cm^2$, respectively.

To further understand the underlying reason for smaller polarization and d_{33} in our material we also calculated the ferroelectric properties for a new composition, $Ba_{0.9}Ca_{0.1}Ti_{0.9}Zr_{0.1}O_3$ (BCTZ3). Note that in this composition we decreased the Zr^{4+} concentration from 0.26 to 0.10. The DFT results are given in Table 1 and the ferroelectric properties for this hypothetical system compare favorably with BCTZ1. Therefore, the inferior d_{33} we observe in BCTZ2 is attributed to the higher concentration of Zr^{4+} cation in the solid solution. Increase in concentration of Zr^{4+} usually induces polar nano-regions, resulting in relaxor-like behavior (42). More importantly, high concentration of dopants can also pin domain wall motion (43). Thus, the overall polarization of the system decreases, as seen in our DFT calculations.

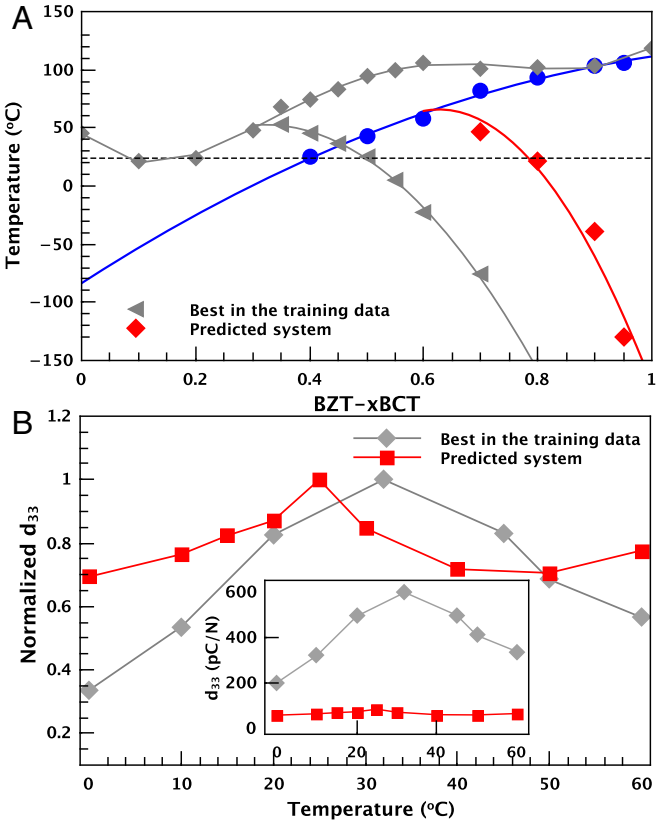

Fig. 4. Results from experiment. (*A*) Phase diagram of newly found BZT-*m*50-*n*30 system, comparing with that of best in the training set. (*B*) The temperature dependence of normalized d_{33} for the newly found BZT-*m*50-*n*30 system, compared with the best in the training set. (*Inset*) Measured d_{33} data as a function of temperature for BZT-*m*50-*n*30 and best in the training set.

Table 1. Comparison of the total energy difference (ΔE, in millielectron volts per formula unit) and spontaneous polarization (ΔP_s, in $\mu C/cm^2$) for the tetragonal (T), orthorhombic (O), and rhombohedral (R) ferroelectric phases from DFT calculations with respect to the paraelectric cubic (C) phase

Materials	$\Delta E^{C\to T}$	$\Delta E^{C\to O}$	$\Delta E^{C\to R}$	ΔP_s^T	ΔP_s^O	ΔP_s^R
BCTZ1	−17	−20	−22	35	38.4	38.2
BCTZ2	−0.6	−0.3	−0.5	0.07	0.05	0.06
BCTZ3	−21	−25	−26	36.8	39.8	39.7

BCTZ1, BCTZ2, and BCTZ3 correspond to $Ba_{0.85}Ca_{0.15}Ti_{0.9}Zr_{0.1}O_3$, $Ba_{0.9}Ca_{0.1}Ti_{0.74}Zr_{0.26}O_3$ (our MPB composition), and $Ba_{0.9}Ca_{0.1}Ti_{0.9}Zr_{0.1}O_3$, respectively.

Discussion

We have shown how to rapidly optimize the composition space in the search for new BCT–BZT and BCT–BTS solid solutions with desired MPBs. Given a vast virtual space of 1,203 solid solutions to choose from for experiments, our strategy found the optimal material in one iteration. We attribute this to the choice of feature sets and domain knowledge that led to the formulation of a predictive model for optimizing the MPB curvature as a function of chemical compositions. As noted earlier, few models exist in the literature for guiding the next experiments for targeted piezoelectric properties. Our work, thus, serves not only as a template for engineering the MPB curvature in other Pb-free piezoelectrics [e.g., $KNbO_3$ solid solutions (44)] but can be applied to optimizing properties of other materials, where physical theories or domain knowledge in the form of functional forms or inequalities exist for predictive modeling. Furthermore, our prediction for the BZT-*m*50-*n*30 compound has led to the investigation of a previously unexplored composition space (in terms of m and n), demonstrating the potential of our approach in the search for new materials that depart from those represented in the training data.

Our training dataset contained both Zr and Sn cations. Our analysis has shown that Zr has a stronger tendency for realizing a vertical MPB relative to the Sn-based solid solutions. In terms of the Shannon's ionic radii (32), Zr^{4+} and Sn^{4+} are 0.72 and 0.69 Å, respectively, with only a 0.03-Å size difference within the hard sphere model. Our DFT calculations for the bulk $BaZrO_3$ and $BaSnO_3$ perovskites in the constrained rhombohedral $R3m$ symmetry also do not show any appreciable difference in the ferroelectric properties (33). We also know from Landau theory that γ_2 in Eq. **1** stabilizes the T phase with respect to the R phase, whereas γ_3 stabilizes the R phase relative to the T phase, thereby influencing the relative tilting affecting the MPB curvature (16). Although d_{33} is generally large in the vicinity of the MPB, and a given degree of tilting can result from free energies with different degrees of anisotropy, the latter are difficult to relate directly. In addition, the separation of the tricritical points about the triple point on the paraelectric-to-ferroelectric transition boundary also influences the magnitude of d_{33}. Thus, more theoretical work and additional experiments are required to clarify the role of dopants on piezoelectric properties.

An important aspect of our work has been to show how a Bayesian approach can be formulated for materials design. Considering the limited experimental data and complex functional relationships among materials descriptors and properties, Bayesian approaches driven by scientific knowledge offer the means to constrain the model space and can lead to robust predictions. Domain knowledge, such as the Landau functional form and the constraints on phase boundaries in this study, helped in navigating the search space by suitable prior construction so that the desired composition could be achieved in a few trials. However, the data-driven approach using widely available regression methods encodes domain knowledge largely in the

feature selection. With the latter method, we used an additional design step with uncertainties to guide the next experiment, and this is important if datasets are small and we need to explore a large unknown search space. Nevertheless, both approaches promise to play a crucial role in realizing a goal of the materials genome initiative (1), which is to minimize the number of experiments required to find materials with targeted properties.

Methods

Ceramic samples were synthesized using a solid-state reaction. Calcining was performed at 1,350 °C and sintering (in air) at 1,450 °C. Phase diagrams were determined from dielectric permittivity (ϵ)–temperature (T) curves and d_{33} was measured using a Berlingcourt-type meter for poled samples. For DFT calculations, core and valence electrons were generated using the Opium code (45) with a plane-wave cutoff for wavefunctions of 60 and kinetic energy 240 Ry. Details are given in *SI Appendix*, Note 6.

ACKNOWLEDGMENTS. We are grateful to John Hogden, James Theiler, and Eli Ben-Naim for discussions. This work was supported by Los Alamos National Laboratory, Laboratory Directed Research and Development program Grant 20140013DR, and Center for Nonlinear Studies (to P.V.B. and T.L.); 973 Program Grant 2012CB619401; National Natural Science Foundation Grants 51302209, 51431007, 51320105014, and 51321003 (to D.X. and R.Y.); and National Science Foundation Grant 1553281 (to T.H., X.Q., and E.R.D.).

1. OSTP (2011) Materials genome initiative for global competitiveness (Office of Science and Technology Policy, Washington, DC). Available at https://www.whitehouse.gov/sites/default/files/microsites/ostp/materials_genome_initiative-final.pdf.
2. Balachandran PV, Broderick SR, Rajan K (2011) Identifying the inorganic gene for high-temperature piezoelectric perovskites through statistical learning. *Proc R Soc A Math Phys Eng Sci* 467(2132):2271–2290.
3. Curtarolo S, et al. (2013) The high-throughput highway to computational materials design. *Nat Mater* 12(3):191–201.
4. Saal J, Kirklin S, Aykol M, Meredig B, Wolverton C (2013) Materials design and discovery with high-throughput density functional theory: The open quantum materials database (OQMD). *JOM* 65(11):1501–1509.
5. Jain A, et al. (2013) Commentary: The materials project: A materials genome approach to accelerating materials innovation. *APL Mater* 1(1):011002.
6. Xue D, et al. (2016) Accelerated search for materials with targeted properties by adaptive design. *Nat Commun* 7:11241.
7. Ghiringhelli LM, Vybiral J, Levchenko SV, Draxl C, Scheffler M (2015) Big data of materials science: Critical role of the descriptor. *Phys Rev Lett* 114(10):105503.
8. Balachandran PV, Theiler J, Rondinelli JM, Lookman T (2015) Materials prediction via classification learning. *Sci Rep* 5:13285.
9. Kim C, Pilania G, Ramprasad R (2016) From organized high-throughput data to phenomenological theory using machine learning: The example of dielectric breakdown. *Chem Mater* 28(5):1304–1311.
10. Dougherty ER, Zollanvari A, Braga-Neto UM (2011) The illusion of distribution-free small-sample classification in genomics. *Curr Genomics* 12(5):333–341.
11. Ruess J, Parise F, Milias-Argeitis A, Khammash M, Lygeros J (2015) Iterative experiment design guides the characterization of a light-inducible gene expression circuit. *Proc Natl Acad Sci USA* 112(26):8148–8153.
12. Devonshire A (1949) XCVI. Theory of barium titanate. *London Edinburgh Dublin Philosophical Mag J Sci* 40(309):1040–1063.
13. Tolédano J, Tolédano P (1987) *The Landau Theory of Phase Transitions* (World Scientific, Singapore).
14. Heitmann AA, Rossetti GA (2014) Thermodynamics of ferroelectric solid solutions with morphotropic phase boundaries. *J Am Ceram Soc* 97(6):1661–1685.
15. Ahart M, et al. (2008) Origin of morphotropic phase boundaries in ferroelectrics. *Nature* 451(7178):545–548.
16. Porta M, Lookman T (2011) Effects of tricritical points and morphotropic phase boundaries on the piezoelectric properties of ferroelectrics. *Phys Rev B* 83(17):174108.
17. Liu W, Ren X (2009) Large piezoelectric effect in Pb-free ceramics. *Phys Rev Lett* 103(25):257602.
18. Xue D, et al. (2011) Elastic, piezoelectric, and dielectric properties of $Ba(Zr_{0.2}Ti_{0.8})O_3$-$50(Ba_{0.7}Ca_{0.3})TiO_3$ Pb-free ceramic at the morphotropic phase boundary. *J Appl Phys* 109(5):054110.
19. Rödel J, et al. (2009) Perspective on the development of lead-free piezoceramics. *J Am Ceram Soc* 92(6):1153–1177.
20. Xue D, et al. (2011) Large piezoelectric effect in Pb-free $Ba(Ti,Sn)O_3$-x(Ba, Ca)TiO_3 ceramics. *Appl Phys Lett* 99(12):122901.
21. Zhou C, et al. (2012) Triple-point-type morphotropic phase boundary based large piezoelectric Pb-free material $Ba(Ti_{0.8}Hf_{0.2})O_3$-$(Ba_{0.7}Ca_{0.3})TiO_3$. *Appl Phys Lett* 100(22):222910.
22. Armiento R, Kozinsky B, Fornari M, Ceder G (2011) Screening for high-performance piezoelectrics using high-throughput density functional theory. *Phys Rev B* 84: 014103.
23. Chelikowsky JR, Phillips JC (1978) Quantum-defect theory of heats of formation and structural transition energies of liquid and solid simple metal alloys and compounds. *Phys Rev B* 17:2453–2477.
24. Saad Y, et al. (2012) Data mining for materials: Computational experiments with *AB* compounds. *Phys Rev B* 85:104104.
25. Pilania G, Gubernatis JE, Lookman T (2015) Structure classification and melting temperature prediction in octet AB solids via machine learning. *Phys Rev B* 91:214302.
26. Seko A, Maekawa T, Tsuda K, Tanaka I (2014) Machine learning with systematic density-functional theory calculations: application to melting temperatures of single- and binary-component solids. *Phys Rev B* 89:054303.
27. Mannodi-Kanakkithodi A, Pilania G, Huan TD, Lookman T, Ramprasad R (2016) Machine learning strategy for accelerated design of polymer dielectrics. *Sci Rep* 6:20952.
28. Seko A, et al. (2015) Prediction of low-thermal-conductivity compounds with first-principles anharmonic lattice-dynamics calculations and Bayesian optimization. *Phys Rev Lett* 115:205901.
29. Balachandran PV, Xue D, Theiler J, Hogden J, Lookman T (2016) Adaptive strategies for materials design using uncertainties. *Sci Rep* 6:19660.
30. Pauling L (1932) The nature of the chemical bond. IV. The energy of single bonds and the relative electronegativity of atoms. *J Am Chem Soc* 54(9):3570–3582.
31. Goldschmidt VM (1926) Die Gesetze der Krystallochemie. *Die Naturwissenschaften* 21:477–485.
32. Shannon RD (1976) Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. *Acta Cryst A* 32:751–767.
33. Balachandran PV, Xue D, Lookman T (2016) Structure–Curie temperature relationships in $BaTiO_3$-based ferroelectric perovskites: anomalous behavior of $(Ba,Cd)TiO_3$ from DFT, statistical inference, and experiments. *Phys Rev B* 93:144111.
34. McQuarrie M, Behnke FW (1954) Structural and dielectric studies in the system (Ba, Ca)(Ti, Zr)O_3. *J Am Ceram Soc* 37(11):539–543.
35. Wei X, Yao X (2007) Preparation, structure and dielectric property of barium stannate titanate ceramics. *Mater Sci Eng B* 137(1):184–188.
36. Abrahams SC, Kurtz SK, Jamieson PB (1968) Atomic displacement relationship to Curie temperature and spontaneous polarization in displacive ferroelectrics. *Phys Rev* 172:551–553.
37. Perdew JP, Burke K, Ernzerhof M (1996) Generalized gradient approximation made simple. *Phys Rev Lett* 77:3865–3868.
38. Giannozzi P, et al. (2009) QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials. *J Phys Condens Matter* 21(39):395502.
39. de Gironcoli S (1992) Phonons in Si-Ge systems: An ab initio interatomic-force-constant approach. *Phys Rev B* 46:2412–2419.
40. Bellaiche L, Vanderbilt D (2000) Virtual crystal approximation revisited: Application to dielectric and piezoelectric properties of perovskites. *Phys Rev B* 61:7877–7882.
41. Bernardini F, Fiorentini V, Vanderbilt D (1997) Spontaneous polarization and piezoelectric constants of III-V nitrides. *Phys Rev B* 56:R10024–R10027.
42. Kleemann W (2014) Relaxor ferroelectrics: Cluster glass ground state via random fields and random bonds. *Phys Status Solidi B* 251(10):1993–2002.
43. Yang T, Gopalan V, Swart P, Mohideen U (1999) Direct observation of pinning and bowing of a single ferroelectric domain wall. *Phys Rev Lett* 82(20):4106.
44. Wang R, et al. (2015) Temperature stability of lead-free niobate piezoceramics with engineered morphotropic phase boundary. *J Am Ceram Soc* 98(7):2177–2182.
45. Opium (2014) Opium – pseudopotential generation project. Available at opium.sourceforge.net.